PROPERTY OF
CETON UNIVERSITY LIBRARY
ECEIVED OCT 14 1940

SW

The PSYCHOLOGICAL RECORD . . . .

SEPTEMBER, 1940
Vol. IV No. 11

ON THE ECONOMICAL ARRANGEMENT OF TOOLS; THE HARMONIC SERIES AND THE PROPERTIES OF SPACE

G. K. ZIPF

THE PRINCIPIA PRESS, INC.
BLOOMINGTON, INDIANA

Price of this number, 35 cents

EDITOR	J. R. KANTOR, *Indiana University*
BUSINESS EDITOR	C. M. LOUTTIT, *Indiana University*

DEPARTMENTAL EDITORS:

ABNORMAL	EDMUND S. CONKLIN, *Indiana University*
CHILD	MARTIN REYMERT, *Mooseheart Laboratory*
CLINICAL	G. A. KELLY, *Fort Hays K. S. C.*
COMPARATIVE	E. A. CULLER, *University of Rochester*
EDUCATIONAL	J. G. PEATMAN, *City College of New York*
EXPERIMENTAL	B. F. SKINNER, *University of Minnesota*
INDUSTRIAL	JOSEPH TIFFIN, *Purdue University*
PHYSIOLOGICAL	C. F. SCOFIELD, *University of Buffalo*
PSYCHOMETRICS	J. P. GUILFORD, *University of Nebraska*
SOCIAL	NORMAN C. MEIER, *University of Iowa*
ASSISTANT EDITOR	J. W. CARTER, JR., *Wichita, Kansas*

The Principia Press, Inc., has undertaken the publication of this co-operative journal to afford authors the opportunity of immediate publication at the least possible expense. The present low cost of publication and possible future reductions, depend entirely upon the number of subscriptions. The subscription price has been made low to induce individuals to subscribe. Under the Articles of Incorporation of the Principia Press no profit can be made on any of its publications. Therefore an increase in the number of subscribers will be reflected in reduced costs to authors and in increase in the number of pages published annually.

EDITORIAL BOARD. The above named board of associate editors have agreed to serve in an active capacity for a period of three years, and all manuscripts will be submitted to them.

MANUSCRIPTS may be sent to the appropriate associate editor or to Dr. J. R. Kantor. Longer papers (30 or more typewritten pages) in the above mentioned fields are especially desired.

COSTS. The cost of publication will be $2.00 per page of approximately 470 words. The charge for line cut illustrations will be $2.00 regardless of size. Special charges will have to be made for complex tabular matter and for half-tone or colored illustrations.

REPRINTS. One hundred copies of each paper will be supplied gratis. Additional copies may be purchased in any quantity at a rate which will be set when the author is informed of the cost of publication.

ROYALTIES. Fifty per cent of the net income from the sale of individual copies, or from copies sold as part of back volumes, will be credited as royalties to the author's account. Royalties cannot be paid on income from subscriptions current in the year of publication.

SUBSCRIPTIONS. The subscription price is $4.00 per year for an annual volume of approximately 500 pages. Individual papers will be sold at an advertised price, which will depend upon the size of the article. Foreign subscriptions will be $4.50.

CORRESPONDENCE concerning editorial matters should be addressed to Dr. J. R. Kantor, Indiana University. Business correspondence should be addressed to Dr. C. M. Louttit, Principia Press, Inc., Bloomington, Ind.

ON THE ECONOMICAL ARRANGEMENT OF TOOLS: THE HARMONIC SERIES AND THE PROPERTIES OF SPACE[1]

By G. K. Zipf,
Harvard University

In the present paper we shall study the underlying principles of the economical arrangement of tools in an ideal workshop where tools are to be transported from their places of repose to the center of the workshop with a minimal expenditure of energy. And we shall see that the harmonic series

$$F\,Sn = \frac{F}{1} + \frac{F}{2} + \frac{F}{3} + \cdots + \frac{F}{n}$$

(where F is the frequency of usage of the most frequently used tool) may be inherent in the arrangement of our tools as defined below. In our demonstration we shall proceed from a consideration of "ideal" conditions to that of the arrangement of actual tools.

A. The Cube of Distance as an Approximate Limitation Upon the Spacing and Number of Possible Tools

If we assume that a tool is a physical mass of material whose use in the performance of a task will save energy, and if we for the present assume—quite arbitrarily—that the more different tools our workshop has, the more efficient its performance will be, then the question arises as to the existence of a possible limitation of any sort upon our tools.

Let us begin our demonstration by assuming that all our tools are equivalent in mass, shape and size; although this assumption oversimplifies our problem, it will nevertheless give us an idea of the limitation upon the number of tools that may be different in mass, shape and size. Furthermore let us assume that gravitational forces are non-operative in our workshop, to the end that we can

[1] Recommended for publication by Dr. J. R. Kantor, August 23, 1940.

Copyright 1940 by G. K. Zipf. All rights reserved.

pack space with tools in all directions. Instead of gravitational forces, we shall imagine that our workshop is filled with a constant friction (μ) which will offer resistance to the movement of our tools. And finally let us assume that an imaginary artisan occupies a fixed point in space around which the tools are arranged as closely to him as possible.

Under the above conditions we may say that the number of possible tools in the workshop will increase directly according to the cube of the distance from the artisan. This will be true because we are packing three dimensional space with discrete and equivalent entities with shortest distances to a central point and hence are dealing with the established properties of spherical volume. This restriction of the cube of the distance from the center of the workshop will lie against the number of tools that can be packed around the artisan. As ever more tools are added to the workshop, then the farther from the artisan the most remote tools will be.

B. The Inverse Square of Distance as a Limitation Upon the Frequency of Transportation of Tools to the Center of the Workshop

In the above section we have arranged tools with an economical use of space with no regard to the minimizing of energy in their transportation to and from the center of the workshop. Let us now inquire into the frequency with which the above-arranged tools will be used.

Since our tools are equivalent entities of matter, then we may say that the amount of energy (E) which will be consumed when transporting the tool through constant friction (μ) to the center of the workshop will be equal to the product of the mass (M) of the tool when multiplied by the total distance (D) that the tool must be transported. In other words:

$$\mu \times M \times D = E$$

The energy (E) would be the artisan's internal energy expended in reaching for and replacing the tool; the physicists would call our energy *work*.

But inasmuch as some tools will be nearer to the artisan than

others, because of the properties of space, then, if the artisan is to conserve energy in transporting them to and from the center of his shop, he must place those tools nearest to him which he will use most often (for we are assuming that he will use all his tools sometimes). Indeed the arrangement of all the various tools in the workshop will depend upon the relative frequency of their use. And we shall assume for the moment that the artisan can redesign tools and change jobs so that he can alter the relative frequency with which any tool is used.

As far as the frequency of usage is concerned, the farther the tool is situated from the artisan, the less often it should be used. And here again the limits upon frequency of usage are quite precise. For, since our tools are equivalent in size, they may be viewed as arranged on concentric spherical shells (s) about the artisan as the center (we refer to a situation of small tools in reference to the radius of the sphere). The actual number (N) of different tools on each spherical shell (s) will increase directly according to the square of the radius of that spherical shell, because that is the established property of the surfaces of spheres. In other words, the number of different tools on a given spherical shell (Ns) will increase according to the square of the distance of that shell to the center.

But since the artisan must minimize the expenditure of energy in using his tools, and since he must use all our equivalent tools, then the frequency with which he uses a given tool will vary inversely with its distance to the center (or $F \times D = \text{Constant}$) —provided that we define the usage of a tool as the transportation of the tool from its place on its spherical shell to the center of the workshop and return (we ignore the use of the tool at the center). Inasmuch as the number of tools in a given shell (Ns) increases according to the square of the shell's distance ($N_s = D^2 \times C$) and since $F \times D = \text{Constant}$, then it follows that

$$N_s F^2 = \text{Constant}$$

This equation says that the number of tools of like frequency of usage will vary inversely according to the square of their frequency of usage,—or *vice versa*.

Nevertheless this equation for number and frequency is not necessarily final because we do not know that the most economical

packing of tools in space is automatically the most economical arrangement of tools to the end that the expenditure of energy will be minimized in using these tools. In other words, we are economizing energy in our workshop; and though we may also accidentally economize space thereby, the presumptions are against it. Let us now arrange our tools with a maximum conservation of energy and note what happens to the equation $N F^2 = \text{Constant}$.

C. The Harmonic Series and the Sum of All Products of $M \times D \times F \times \mu$ in Three Dimensional Space

Since we are assuming that the usage of a given tool a single time will consume energy directly in proportion to its total distance to and from the center of the workshop, and since we are assuming that all tools are equivalent in mass, shape and size, then it follows that the frequency (F) with which any tool is used will be inversely proportionate to its distance from the center of the workshop. This is true first because the energy (E) consumed in the usage of any tool is governed by the equation, $\mu \times M \times D = E$, and second because we have assumed that all tools have the same M, and third because μ represents a constant friction throughout the workshop.

In short, if we choose to conserve energy in our workshop, the equation, $D \times F = \text{Constant}$, will apply to the frequency of usage of every tool in the workshop. As the distance of the tool from the center increases, its frequency of usage will decrease proportionately.

But we also know that as the distance increases from the center of the workshop, the maximum possible number (N_s) of tools of that distance will increase according to the square of that distance, and hence that the frequency of usage (F) of that number (N_s) will decrease according to the inverse square (i.e. $N_s F^2 = \text{Constant}$). Let us now seek to determine what will be the number of tools (N_F) of like frequency of usage (F) under the condition that $D \times F = \text{Constant}$ throughout the entire workshop.

Let us begin by defining as a *unit usage* the round-trip of a tool from its resting place to the center of the workshop and return.

Furthermore let us define as a *unit interval of measurement* the time necessary to use all the tools in the shop according to the equation, D x F = Constant, under the provision that the most distant tools will have been used only once. And finally let us assume that there are no fractional frequencies of usage. In short, during our unit interval of measurement, all our tools will have been used an integral number of times which will be inversely proportionate to their distance from the center, and yet the most distant tool or tools will have been used only once.

Now at the end of the unit interval of measurement, all those tools that have been used the same number of times (i.e. that have had a like frequency of usage) will be equidistant from the center and hence be situated on the same spherical shell. In short, all tools of like frequency of usage will belong together as exclusive members of their respective spherical shell, and their number may be represented by N_F (in which F represents the like frequency of usage of each of the N tools).

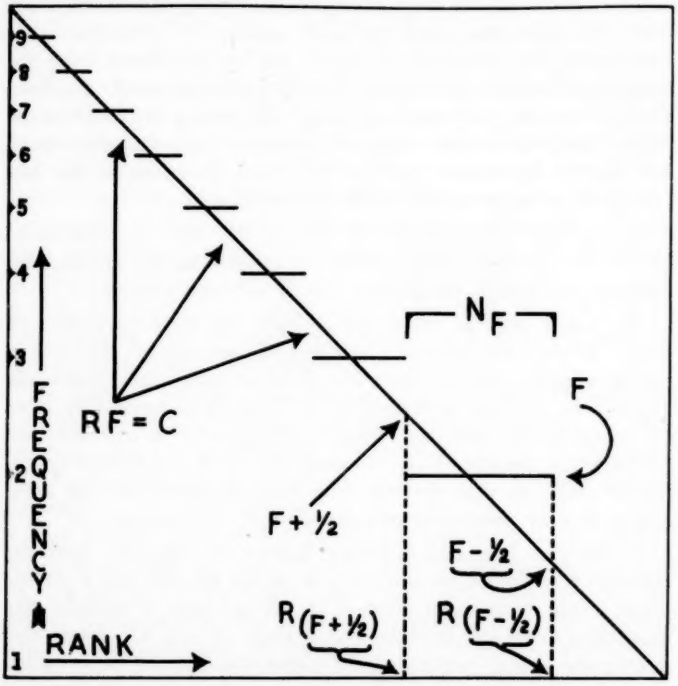
Let us now try to derive the size of N_F from our equation, D x F = Constant, and thereby try to express the minimal expenditure of energy (as defined above) in terms of space.

We know that all tools cannot have the same frequency of usage. If after the lapse of a unit interval we rank all our tools in the decreasing order of their frequency of usage, then the rank (R) of any tool will be equivalent to the tool's distance (D) from the center of the workshop, because the size of F is inversely proportionate to the size of D. In short, R = D. And granting the proper units, we may say that R x F = Constant for the same reason that we previously said that D x F = Constant.

If we plot (see accompanying figure) on log. log. grid the equation R x F = Constant, with R on the abscissa and F on the ordinate for frequencies from 1 to 10, we have a straight line descending from left to right with a slope of —1. Of course this is only the lowest portion of the curve which could be extended upward and to the left indefinitely, depending upon the number of different tools in the workshop, since we are specifying only the frequencies of occurrence and not the ranks.

Now the points on the line, R x F = C, represent a complete continuity of values for R and for F. Since, by definition, all fre-

quencies (F) must be integral, then the points on the line, R x F = C, which lie between the respective frequencies (F) of 1, 2, 3, 4 and so on, do not represent the actual frequencies of usages of actual tools. The actual frequencies of usages of actual tools are indicated by the horizontal lines (or *treads*) that are drawn to cut the diagonal line at the integral frequencies. Thus, for example, the *tread* marked N_F at the frequency of 2 represents the number of different tools, which, during a unit interval, are used twice.



Let us now describe the criteria used in drawing the tread N^F; for, since the same criteria were used for all other treads, by describing the tread, N_F, we shall automatically describe all other treads, *mutatis mutandis*.

The tread, N_F, (with $F = 2$ arbitrarily), is theoretically the projection on $F = 2$ of that portion of the line, $RF = C$, which falls between $F + \frac{1}{2}$ and $F - \frac{1}{2}$, thus marked. That is, since we may have only an integral number of frequencies, we should expect for each frequency-integral a number of ranks (i.e. a number of different tools) represented by the ranks (R) on $RF = C$ halfway between the integral frequency above, on the one hand, and the integral frequency below, on the other.

As for the size of N_F (i.e. the number of different tools which are used twice during a unit interval), we can estimate it by subtracting the rank of the tool at $F + \frac{1}{2}$ from the rank of the tool at $F - \frac{1}{2}$; in short, by subtracting $R_{(F+\frac{1}{2})}$ from $R_{(F-\frac{1}{2})}$ we shall have N_F. Let us proceed with this subtraction.

Working with our equation, $RF = C$, we know the following facts about $R_{(F-\frac{1}{2})}$ and $R_{(F+\frac{1}{2})}$:

$$R_{(F-\frac{1}{2})} = \frac{C}{F-\frac{1}{2}} \text{ and } R_{(F+\frac{1}{2})} = \frac{C}{F+\frac{1}{2}}$$

hence, we may say: $N_F = \frac{C}{F-\frac{1}{2}} - \frac{C}{F+\frac{1}{2}}$

which, after clearing fractions, simplifying and transposing becomes

$$N_F \ (F^2 - \tfrac{1}{4}) = C$$

In other words, the number of tools of like frequency of usage during a unit interval, namely N_F, when multiplied by the square of their frequency minus $\frac{1}{4}$, remains constant.

If we may say that the equation, $NF^2 = C$, represents the most economical frequency of usage of our defined tools *after* they have been arranged with maximum economy of space; and if we may say that the equation, $N_F \ (F^2 - \frac{1}{4}) = C$, represents the most economical usage of our defined tools in space, then we may say that the most economical usage of our defined tools in space is not the most economical use of space. In short, it would seem that *in minimizing energy we waste space*. The constant fraction, $\frac{1}{4}$, in the last preceding formula represents the amount of space we waste, in our model workshop, and we shall formally define it as the *spacial leak*.

But now that we have inspected the conditions at the bottom of our line, $RF = C$, let us inspect the conditions at the top—

all in terms of all our preceding definitions and postulates. (As we increase the size of F, the difference between F^2 and $F^2 - \frac{1}{4}$ becomes very small. Thus, if $F = 100$, then $F^2 = 10{,}000$, and $F^2 - \frac{1}{4} = 9{,}999\frac{3}{4}$.)

If maximum economy of frequency of movement of tools over distance is represented by $R_t F = C$, where R_t represents the rank of a given tool, t, when arranged in decreasing order of frequency of usage, then R_t must be an integral number, because we are assuming that our tools are discrete entities (i.e. there are no fractional tools).

Let us assume that there is a tool of which we may say that $R = 1$. In that case we may say that its frequency of usage will be 1. Furthermore, if there are tools of which we may say that $R = 2$, $R = 3$, $R = 4$ and so on down the integers to n, then we may say that there will be corresponding frequencies of $\frac{1}{2}, \frac{1}{3}, \frac{1}{4} \ldots \frac{1}{n}$. In short, F, which varies inversely with the size of R, will fall into frequency-categories according to the harmonic series $\frac{F}{1}, \frac{F}{2}, \frac{F}{3}, \frac{F}{4} \ldots \frac{F}{n}$. In fact, if each tool has a different frequency of usage, and each has an integral rank, with no integral rank lacking between 1 and n, then we may say that the sum of all our frequencies in decreasing order of size during a unit interval, under the above conditions, could be represented as

$$F\,Sn = \frac{F}{1} + \frac{F}{2} + \frac{F}{3} + \cdots + \frac{F}{n}$$

in which Sn equals the sum of the first n terms of the series and in which F is the frequency of usage of the tool, $R = 1$.

The right-hand member of the above equation represents the sum of the proportions of a harmonic series, with the constant multiplier, F. And we have seen that these proportions may be the natural consequence of the economical use of tools in three dimensional space, according to our previous definitions and postulates.

D. Actual Conditions in Which the Sum of All Products of M x D x F x μ is a Minimum

Up to this point we have dealt with oversimplified conditions. Thus, for example, we have assumed that all tools were equivalent

in mass, size and shape which will hardly be the case in any actual workshop. Furthermore we have tacitly assumed that the artisan could redesign tools and alter jobs so that he could use tools with frequencies that would be inversely proportionate to their respective distances, and so on.

Nevertheless our formulation of ideal conditions is of possible value in indicating the possible goal of an economical arrangement of tools in three dimensional space for the following reasons. After all, the amount of energy (E) necessary to move a mass (M) over a given distance (D) is equal to the product of M x D, if all else is equal. Thus, with a constant distance, the larger the mass, the greater the energy that will be necessary to move it; hence it is economical to decrease the mass of a tool if one can do so without altering its utility. Furthermore, with a constant mass, the longer the distance, the greater the energy that will be necessary to move the mass over that distance; hence it is economical to decrease the distance of the tool from the center of the workshop. In other words it is economical to devise new tools and new arrangements of tools so that the total sum of M x D for all tools will decrease.

But that is not all. The average rate of usage of a given tool (F) should decrease as the size of its product, M x D, increases, and vice versa. In short, economy can be effected by reducing the sum of the products of M x D x F for all tools in the workshop. Similarly the amount of friction, μ, should be reduced, as well as all other factors that represent a consumption of energy in the transportation of tools in the workshop.

Indeed, if an artisan seeks to arrange and use his tools with a minimal expenditure of energy, he will seek to minimize the total of all products of M x D x F x μ for the tools of his workshop. And in this connection we must not overlook the possibility of saving energy (under an assumed constancy of jobs to be performed by the tools) by decreasing the number of tools in the shop (i.e. by devising tools that will take over the tasks of other tools and hence obviate them). But whether or not any new tool should be used, or any old tool revised will depend upon its effect upon the total consumption of energy in terms of some unit interval of measurement (if we assume a constancy of jobs to be performed).

In short, the *comparative value* of any tool in respect to an alternate tool in the performance of a given task in a given situa-

tion will depend entirely upon the amount of energy saved by using the one tool instead of the other in the performance of a task where either tool will be adequately serviceable. That tool will have the greater comparative value whose usage will save the greater total amount of energy.

Up to this point we have said little of the tasks to be performed by the tools of our workshop. If we assume that our workshop has fixed jobs to perform, then the artisan can only seek tools of greater comparative value. But if we assume the presence of the workshop with the tools, then the artisan must seek jobs to perform that will minimize the expenditure of energy in using the tools. Thus if an artisan possesses a carpenter's shop, he must seek jobs for his carpenter's tools, for those jobs will be presumably the most economical ones for his tools. But if he has a specific job to perform, then he must seek those tools which will perform his job most economically. In either case, the artisan seeks to make the total of all products of M x D x F x μ a minimum. Furthermore the same problem of the minimal total will confront the artisan who seeks to keep his workshop busy by adapting tools to processes and processes to tools according to the exigencies of a changing economic system (to be treated in greater detail in a separate publication entitled *The Laws of National Forces*).

Thus far we have made little mention of the size of tools. Since one can pack a larger number of smaller tools nearer the artisan, it is economical to decrease the size (S) of a tool (and we might add S to our factors) provided that a decrease in size does not consume more energy in the total handling of the tool.

But when we speak of the total handling of a tool we mean not only its transportation to and from the center of the workshop but also its use at the center of the workshop. Nevertheless, if we assume that, in using the tool at the center of the workshop, the artisan is expending energy in moving masses over distances, then he will at all times be confronted by the same general problem of minimizing the total of all the products of M x D for all tools.

In short, if we have discrete tools, and if we will minimize the expenditure of energy, then we should seem to encounter—as a limit to economy—the proportions of a harmonic series as an inevitable consequence of the properties of three dimensional space.

The proportions of the harmonic series will also appear in two dimensions. As a homely example let us take a penny and put a ring of pennies around it so that each penny will touch the central penny and both of its neighbors. There will be six pennies in this first ring; twelve in the second; eighteen in the third and so on. The number (N) of the pennies of each successive ring will increase in proportion to the distance ($\frac{N}{D}$ = Constant). The "frequency of usage" of each "penny" in two dimensions will be NF = Constant. Yet let us note the nature of the frequency-categories under our above assumption of discrete equivalent entities in two dimensions. Since the distances of the pennies to the center will fall into the classes of 1, 2, 3, 4 . . . n units, then the frequencies of usages (being inversely proportionate to the distances) will fall into the classes of: $\frac{F}{1}, \frac{F}{2}, \frac{F}{3}, \frac{F}{4} \ldots \frac{F}{n}$, in which $\frac{F}{1}$ represents the highest frequency of usage (the frequency of usage of each of the six pennies in the first ring). As to the stationary penny in the center—that is the artisan.

Hence it would seem to the author that the assumption of a minimal expenditure of energy in the movement of discrete masses of materials over distance in reference to a central point will inevitably entail the proportions of the harmonic series as a limit of economical organization. Furthermore an artisan who is bent upon devising and arranging tools in his workshop for his economical handling, would seem to be confronted by the problem of minimizing the total of all products of M x D x F x μ with the inclusion of the additional factors, such as size, gravitation and the like.

E. Empiric Data

The author has presented the above analysis of the possible relationship between the proportions of a harmonic series and the dimensions of space for a very obvious reason. For he has observed many instances of the proportions of the harmonic series in sets of objective data. For example, the frequencies of occurrences of words in the stream of speech can be found to approximate the formula:

$$F\ Sn = \frac{F}{1} + \frac{F}{2} + \frac{F}{3} + \ldots + \frac{F}{n}$$

in which $\frac{F}{1}$ equals the frequency of occurrence of the most frequently used word (cp. G. K. Zipf, *Psycho-Biology of Language*, Boston, II ed. 1939, and also G. K. Zipf, "Homogeneity and Heterogeneity in Language," *Psychological Record*, II [1938], 347-367). Furthermore in these data we find in fact $N\ (F^2 - \frac{1}{4}) = \text{Constant}$.

Moreover if we set the generalized equation:

$$A\ Sn = \frac{A}{1^p} + \frac{A}{2^p} + \frac{A}{3^p} + \cdots + \frac{A}{n^p}$$

(cp. G. K. Zipf, "The Generalized Harmonic series as a fundamental principle of social organization," *The Psychological Record*, IV [1940] p. 43) we find, for example, that the sizes of communities in the United States and in other countries approximate this series very closely, with $\frac{A}{1}$ equal to the size of the largest community and with values of P lying between 0 and 1 inclusive (to be reported in detail in the author's study, *The Laws of National Forces*).

Furthermore the author has been informed that the amounts of carbon, oxygen and nitrogen in a protein molecule—despite its great complexity and despite variations in structure—occur in amounts of about 51%, 25% and 16% respectively. And here the author observes that these approximate amounts are approximately the proportions of the series 1, $\frac{1}{2}$, $\frac{1}{3}$. Indeed we may represent approximately the protein molecule as

$$M\ Sn + (\text{Residue}) = \frac{M}{1} + \frac{M}{2} + \frac{M}{3} + (\text{Residue})$$

in which M = the mass of carbon.

The author believes that there must be some natural reason for the appearance of these orderly distributions. Indeed he has suspected from the beginning that a law of economy of effort lies behind these unmistakable manifestations of the proportions of a harmonic series. Not only has he attempted to explain the reason for the community-distribution in terms of mass-production and the specialization of labor (ibid.); but he has also tried in the present paper to discuss these harmonic proportions in terms of the economy of energy in the handling of materials in space.

And yet our task has only begun—for, why does man bother to have a workshop?

F. The Hypothesis of the Generalized Harmonic Series

The author has come to believe—as a working hypothesis—that the generalized harmonic series

$$A\,Sn = \frac{A}{1^t} + \frac{A}{2^t} + \frac{A}{3^t} + \cdots + \frac{A}{n^t}$$

may be the principle of much other organization. He believes (subject to the correction of future findings) that the values of t from $-\infty$ through o may well represent the organization of stellar bodies, with t representing time and passing from $-\infty$ to 0. With this hypothesis the author hopes perhaps to explain the equation of Mass-Luminosity [i.e. $M = f(L)$].

Furthermore the author believes that the values of t from 0 to 1 inclusive represent biological organization. And finally he believes that the values of t from 1 to $+\infty$ represent mental organization. In this last connection he here reports the observation of a frequency-distribution of the words in the letters of a patient suffering from paranoia; in this frequency-distribution the size of t is greater than 1.

This working hypothesis which the author reserves for himself, together with all other problems in this present paper, will be treated in detail in a subsequent publication in which supporting data will be advanced. In this subsequent publication, he will attempt to show that the *physical,* the *biological* and the *psychological* are three different types, or phases, of organization of one substance. And in that subsequent publication he will seek both to show that the organism may be a "workshop," and to answer the question: why the "workshop"?